



Figure 1. N-linked oligosaccharides for glycosylation

Peptide sequence (SEQ ID NO:5):

MPFVNKQFNY KDPVNGVDIA YIKIPNAGQM QPVKAFKIHN KIWVIPERDT FTNPEEGDLN  
 PPPEAKQVPV SYDSTYLST DNEKDNLYKG VTKLFERIYS TDLGRMLLTS IVRGIPFWGG  
 STIDTELKVI DTNCINVIQP DGSYRSEELN LVIIGPSADI IQFECKSFGH EVLNLTRNGY  
 GSTQYIRFSP DFTFGFEESL EVDTNPLLGA GKFATDPAVT LAHELIHAGH RLYGIAINPN  
 RVFKVNTNAY YEMSGLEVSE EELRTFGGHD AKFIDSLQEN EFRLYYYNKF KDIASLTNKA  
 KSIVGTTASL QYMKNVFKEK YLLSEDTSKG FSVDKLFKFDK LYKMLTEIYT EDNFVKFFKV  
 LNRKTYLNFD KAVFKINIVP KVNYYTYDGF NLRNTNLAAN FNGQNTTEINN MNFTKLKNFT  
 GLFEFYKLLC VRGIITSKTK SLDKGYNKAL NDLCIKVNNW DLFFSPSEDN FTNDLNKGEE  
 ITSNTNIEAA EENISLDLIQ QYYLTFFNFDN EPENISLENL SSDIIGQLEL MPNIERFPNG  
 KKYELDQYTM FHYLRAQEFH HGKSRIALTN SVNEALLNPS RVYTFSSDY VKKVNKATEA  
 AMFLGWVEQL VYDFTDETSE VSTTDKIADI TIIIPYIGPA LNIGNMPLYD DFGALIFSG  
 AVILLEFIPE IAIPVLGTFA LVSYIANKVL TVQTIDNALS KRNEKWDEVY KYIVTNWLAK  
 VNTQIDLIRK KMKEALENQA EATKAIINYQ YNQYTEEEKN NINFNIDDL SKLNESINKA  
 MININKFLNQ CSVSYLMNSM IPYGVKRLD FDASLKDALL KYIYDNRGTL IGQVDRKLDK  
 VNNTLSTDIP FQLSKYVDNQ RLLSTFTEYI KNIINTSILN LRYESNHLID LSRYSKINI  
 GSKVNFDPID KNQIQLFNLE SSKIEVILKN AIVYNSMYEN FSTFSWIRIP KYFNSISLNN  
 EYTIINCMEN NSGKWVSLNY GEIWTLQDT QEIKQRVVFK YSQMINISDY INRWIFVTIT  
 NNRLNNSKIY INGRLLDQKP ISNLGNIHAS NNIMFKLDGC RDTHRYIWIY YFNLFDKELN  
 EKEIKDLYDN QSNNGILKDF WGDYLYQYDKP YMLNLYDPN KYVDVNNVGI RGYMYLKGPR  
 GSVMTTNIYL NSSLYRGTFE IIKKYASGNK DNIVRNNDRV YINVVVKNE YRLATNASQA  
 GVEKILSALE IPDVGNLSQV VVMKSKNDQG ITNKCKMNLQ DNNGNDIGFI GFHQFNNAIK  
 LVASNWNRYQ IERSSRTLGC SWEFIPVDDG WGERPL

Peptides containing the motif 'N-X-S/T/C (X not P)'(underlined):

position	Peptide	SEQ ID NO:
167-177	SFGHEVL <u>NL</u> TR	6
382-393	V <u>NY</u> TYDGFNL	7
394-415	NTNLAANFNGQNTTEINN <u>MN</u> FTK	8
418-427	<u>NFT</u> GLFEFYK	9
457-477	VNNWDLFFSPSEDN <u>FT</u> NDLN K	10
478-536	GEEITSNTNIEAAEENISLD LIQQYYLTFFNFDNEP <u>EN</u> IS	11
773-779	<u>LN</u> ESINK	12
787-806	FLNQCSVSYLMNSMIPYGVK	13
841-855	VNNTLSTDIPFQLSK	14
872-882	NIINTSILNLR	15
930-948	NAIVYNSMYEN <u>FS</u> TSFWIR	16
952-975	YFNSISLNNNEYTIINCMENN <u>SG</u> WK	17
1001-1013	YSQMINISDYINR	18
1024-1028	<u>LN</u> NSK	19
1086-1098	DLYDNQSNNGILK	20
1141-1156	GSVMTTNIYLNSSLYR	21
1193-1204	LATNASQAGVEK	22
1205-1224	ILSALEIPDVGNLSQVVVMK	23

Figure 1 Continue.

Peptide sequence (SEQ ID NO:39):

KTKSLDKGYN KALNDLCIKV NNWDLFFSPS EDNFTNDLNK GEEITSDTNI EAAEENISLD  
 LIQQYYLTFN FDNEPENISI ENLSSDIIGQ LELMPNIERF PNGKKYELDK YTMFHYLRAQ  
 EFEHGKSRIA LTNSVNEALL NPSRVYTFFS SDYVKKVNKA TEAAMFLGWV EQLVYDFTDE  
 TSEVSTTDKI ADITIIIPYI GPALNIGNML YKDDFVGALI FSGAVILLEF IPEIAIPVLG  
 TFALVSYIAN KVLTVQTIDN ALSKRNEKWD EVYKYIVTNW LAKVNTQIDL IRKKMKEALE  
 NQAEATKAI NYQYNQYTEE EKNNINFNID DLSSKLNESI NKAMININKF LNQC SVSYLM  
 NSMIPYGVKR LEDFDASLKD ALLKYIYDNR GTLIGQVDRL KDKVNNTLST DIPFQLSKYV  
 DNQRLSTFT EYIKNIINTS ILNLRYESNH LIDLSRYASK INIGSKVNFD PIDKNQIQLF  
 NLESSKIEVI LKNAIVYNSM YENFSTSFWI RIPKYFNSIS LNNEYTIINC MENNSGWKVS  
 LNYGEIIWTL QDTQEIKQRV VFKYSQMINI SDYINRWIFV TITNNRLNNS KIYINGRLID  
 QKPISNLGNI HASNNIMFKL DGCRDTHRYI WIKYFNLFDK ELNEKEIKDL YDNQSN SGIL  
 KDFWGDYLQY DKPYMYMLNLY DPNKYVDVNN VGIRGYMYLK GPRGSVMTTN IYLNSSLYRG  
 TKFIIKKYAS GNKDNIVRNN DRVYINVVVK NKEYRLATNA SQAGVEKILS ALEIPDVGNL  
 SQVVVMKSKN DQGITNKCKM NLQDNNGNDI GFIFGHQFNN IAKLVASNWY NRQIERSSRT  
 LGCSWEFIPV DDGWERPL

Peptides containing the motif 'N-X-S/T/C (X not P)':

position	peptide	SEQ ID NO:
20-40	VNNWDLFFSPSEDNFTNDLN K	25
41-99	GEEITSDTNIEAAEENISLD LIQQYYLTFNFDNEPENISI	26
336-342	ENLSSDIIGQLELMPNIER	27
350-369	LNESINK	28
404-418	FLNQCSVSYLMNSMIPYGVK	29
435-445	VNNTLSTDIPFQLSK	30
493-511	NIINTSILNLR	31
515-538	NAIVYNSMYENFSTSFWIR	32
564-576	YFNSISLNNEYTIINCMENN SGWK	33
587-591	YSQMINISDYINR	34
649-661	LNN SK	35
704-719	DLYDNQSN SGILK	36
756-767	GSVMTTNIYLNSSLYR	37
768-787	LATNASQAGVEK	38
	ILSALEIPDVGNLSQVVVMK	

Figure 2.

Peptide sequence (SEQ ID NO: 39):

MPFVNKQFNY KDPVNGVDIA YIKIPNAGQM QPVKAFKIHN KIWVIPERDT FTNPEEGDLN  
 PPPEAKQVPV SYDSTYLST DNEKDNYLKG VTKLFERIYS TDLGRMLLTS IVRGIPFWGG  
 STIDTELKVI DTNCINVIQP DGSYRSEELN LVIIGPSADI IQFECKSFGH EVLNLTRNGY  
 GSTQYIRFSP DFTFGFEESL EVDTNPLLGA GKFDTPAVT LAHELIIHAGH RLYGIAINPN  
 RVFKVNTNAY YEMSGLEVSF EELRTFGGHD AKFIDSLQEN EFRLYYYNKF KDIASLTNKA  
 KSIVGTTASL QYMKNVFKEK YLLSEDTSKG FSVDKLFKFDK LYKMLTEIYT EDNFVKFFKV  
 LNRKTYLNFD KAVFKINIVP KVNYTIYDGF NLRNTNLAAN FNGQNTTEINN MNFTKLKNFT  
 GLFEFYKLLC VRGIITSKTK SLDKGYNKAL NDLCIKVNNW DLFFSPSEDN FTNDLNLKGEE  
 ITSNTNIEAA EENISLDLIQ QYYLTFNFDN EPENISIEENL SSDIIGQLEL MPNIERFPNG  
 KKYELDKYTM FHYLRAQEFH HGKSRIALTN SVNEALLNPS RVYTFSSDY VKKVNKATEA  
 AMFLGWVEQL VYDFTDETSE VSTTDKIADI TIIIPYIGPA LNIGNMLYKD DFVGALIFSG  
 AVILLEFIPE IAIPVLGTFA LVSYIANKVL TVQTIDNALS KRNEKWDEVY KYIVTNWLAK  
 VNTQIDLRK KMKEALENQA EATKAIINYQ YNQYTEEEKN NINFNIDDL SKLNESINKA  
 MININKFLNQ CSVSYLMNSM IPYGVKRLD FDASLKDALL KYIYDNRGTL IGQVDRKDK  
 VNNTLSTDIP FQLSKYVDNQ RLLSTFTEYI KNIINTSILN LRYESNHLID LSRYASKINI  
 GSKVNFDPID KNQIQLFNLE SSKIEVILKN AIVYNSMYEN FSTSFWRIP KYFNSISLNN  
 EYTIINCMEN NSGWKVSILN GEIITWLQDT QEIKQRVVFK YSQMINISDY INRWIFVTIT  
 NNRLNNSKIY INGRLLDQKP ISNLGNIHAS NNIMFKLDGC RDTHRYIWIW YFNLFDEKLN  
 EKEIKDLYDN QSNIGLIDKF WGDYLYQYDKP YMLNLYDPN KYVDVNNVGI RGYMYLKGPR  
 GSVMTTNIYL NSSLYRGTKF IIKKYASGNK DNIVRNNDRV YINVVVKNE YRLATNASQA  
 GVEKILSALE IPDVGNSQV VVMKSKNDQG ITNKCKMNLQ DNNGNDIGFI GFHQFNNAK  
 LVASNWNRY IERSSRTLGC SWEFIPVDDG WGERPL

Peptides containing S or T (underlined):

position	peptide	SEQ ID NO:
49-66	DTFTNPEEGDLNPPPEAK	40
67-84	QVPVSYYDSTYLSTDNEK	41
90-93	GVTK	42
98-105	IYSTDLGR	43
106-113	MLLTSIVR	44
114-128	GIPFWGGSTIDTELK	45
129-145	VIDTNCINVIQPDGSR	46
146-166	SEELNLVIIGPSADIIQFEC K	47
167-177	SFGHEVLNLTR	48
178-187	NGYGSTQYIR	49
188-212	FSPDFTFGFEESLEVDTNPL LGAGK	50
213-231	FATDPAVTLAHELIIHAGHR	51
245-264	VNTNAYYEMSGLEVSFEELR	52
265-272	TFGGHDAK	53
273-283	FIDSLQENEFR	54
292-299	DIASLTNK	55
302-314	SIVGTTASLQYMK	56
321-330	YLLSEDTSKG	57
331-335	FSVDK	58
344-356	MLTEIYTEDNFVK	59
365-371	TYLNFDK	60
382-393	VNYTIYDGFNLR	61
394-415	NTNLAANFNGQNTTEINNMNF TK	62
418-427	NFTGLFEFYK	63
433-438	GIITSK	64
439-440	TK	
441-444	SLDK	66

Figure 2

Continue

457-477	VNNWDLFF <u>S</u> P <u>S</u> EDNFTNDLN K	67
478-536	GEEITSDT <u>N</u> IEAAEENISLD LIQQYYLT <u>F</u> NFDNEPENIS <u>I</u>	68
548-555	ENL <u>S</u> SDIIGQLELMPNIER	69
564-565	YTMFH <u>Y</u> LR	
566-581	<u>S</u> R	
582-592	IALTNSVNEALLN <u>P</u> SR	71
597-626	VYTF <u>F</u> SSDYVK	72
627-649	ATEAAMFLGWVEQLVYDFTD ETSEVSTTDK	73
650-688	IADIT <u>I</u> IIPYIGPALNIGNM LYK	74
689-701	DDFVGALIFSGAVILLEFIP EIAIPVLGT <u>F</u> FALV <u>S</u> YIANK	75
712-720	VLTVQTIDNAL <u>S</u> K	76
721-729	YIVT <u>N</u> WLAK	77
734-744	VNTQIDLR	78
745-759	EALNQAEAT <u>K</u>	79
760-772	AIINYQYNQYTEEEK	80
773-779	NNINFNIDDL <u>S</u> SK	81
787-806	LNE <u>S</u> INK	82
808-816	FLNQCSVS <sup>Y</sup> LMNSMIPYGVK	83
828-836	LEDFDAS <u>L</u> K	84
841-855	GTLIGQVDR	85
862-871	VNNTLSTDIPFQL <u>S</u> K	86
872-882	LLSTFTEYIK	87
883-893	NIINT <u>S</u> SILNLR	88
894-897	YESNHLIDL <u>S</u> R	89
903-908	YAS <u>K</u>	90
912-923	INIG <u>S</u> K	91
930-948	NQIQLFNLE <u>S</u> SK	92
952-975	NAIVYNSMYENFSTSFWIR	93
976-994	YFNSISLNNEYTIINCMENN <u>S</u> GWK	94
1001-1013	VSLNYGEIIWTLQDTQEIK	95
1014-1023	YSQMINISDYINR	96
1024-1028	WIFVTITNNR	97
1035-1056	LNN <u>S</u> K	98
1062-1065	LIDQKPISNLGNIHASNNIM FK	99
1066-1086	DTHR	100
1087-1098	DLYDNQSN <u>S</u> GILK	101
1101-1116	G <u>S</u> VMTTNIYLN <u>S</u> SLYR	102
1117-1159	GTK	
1160-1170	YASGNK	104
1171-1204	LATNASQAGVEK	103
1205-1224	ILSALEIPDVGNL <u>S</u> QVVVMK	70
1225-1226	SK	
1227-1234	NDQGITNK	65
1235-1269	LVA <u>S</u> NWYNR	24
1270-1276	<u>S</u> SR	
1277-1296	TLGCSWEFIPVDDGWGERPL	105